

Amendments to the Specification

Please replace the paragraph beginning at line 9 on page 5 with the following amended paragraph:

Using the method of Benight et al. a family of 100 sequences was obtained using a computer algorithm to have optimal hybridization properties for use in nucleic acid detection assays. The sequence set of 100 oligonucleotides was characterized in hybridization assays, demonstrating the ability of family members to correctly hybridize to their complementary sequences with an absence of cross hybridization. These are the sequences having SEQ ID NOS: ~~1 to 100~~ 1173 to 1272 of Table I. This set of sequences has been expanded to include an additional 110 sequences that can be grouped with the original 100 sequences as having non-cross hybridizing properties, based on the characteristics of the original set of 100 sequences. These additional sequences are identified as SEQ ID NOS:~~101 to 210~~ 1273 to 1382 of the sequences in Table I. How these sequences were obtained is described below.

Please replace the paragraph beginning at line 10 on page 12 with the following amended paragraph:

A phantom sequence may thus be generated from exemplary Sequence 1 and Sequence 2 as follows:

Sequence 1: ATGTTTAGTGAAAGTTAGTATTG (SEQ ID NO:1383)

* •

Sequence 2: ATGTTAGTGAATAGTATAGTATTG (SEQ ID NO:1384)

• ♦

Phantom Sequence: ATGTTAGTGAAAGTTAGTATTG

Please replace the paragraph beginning at line 32 on page 53 with the following amended paragraph:

A preferred family of 100 tags is shown as SEQ ID NOS:~~1 to 100~~ 1173 to 1272 in Table I. Characterization of the family of 100 sequence tags was performed to determine the ability of these sequences to form specific duplex structures with their complementary sequences and to assess the potential for cross hybridization. The 100 sequences were synthesized and spotted onto glass slides where they were coupled to the surface by amine linkage. Complementary tag sequences were Cy3-labeled and hybridized individually to the array containing the family of 100 sequence tags. Formation of duplex structures was detected and quantified for each of the positions on the array. Each of the tag sequences performed as expected, that is the perfect match duplex was formed in the absence of significant cross hybridization under stringent hybridization conditions. The results of a sample hybridization are shown in Figure 1. Figure 1a shows the hybridization pattern seen when a microarray containing all 100 probes was hybridized with the target complementary to probe 181234. The 4 sets of paired spots correspond to the probe complementary to the target. Figure 1b shows the pattern seen when a similar array was hybridized with a mix of all 100 targets. These results indicate that the family of sequences which is the subject of this patent can be used as a family of non-cross hybridizing (tag) sequences.

Please replace the paragraph beginning at line 16 on page 54 with the following amended paragraph:

The family of 100 non-cross-hybridizing sequences can be expanded by incorporating additional tetramer sequences that are used in constructing further 24mer oligonucleotides. In one example, four additional words were included in the generation of new sequences to be considered for inclusion as non-cross talkers in a family of sequences that were obtained from the above method using 10 tetramers. In this case, the four additional words were selected to avoid potential homologies with all potential combinations of other words: YYXW (TTAG); WYYX (GTTA); XYYW (ATAG) and WYYY (GTTT). The total number of sequences containing six words using the 14 possible words is 14^6 or 7,529,536. These sequences were screened to eliminate sequences that contain repetitive regions that present potential hybridization problems such as four or more of a similar base (e.g., AAAA or TTTT) or pairs of G's. Each of these sequences was compared to the sequence set of the original family of 100 non-cross-hybridizing sequences (SEQ ID NOs:~~1 to 100~~ 1173 to 1272). Any new sequence that contained a minimal threshold of homology (that does not include the use of insertions or deletions) such as 15 or more matches with any of the original family of sequences was eliminated. In other words, if it was possible to align a new sequence with one or more of the original 100 sequences so as to obtain a maximum simple homology of 15/24 or more, the new sequence was dropped. Simple homology" between a pair of sequences is defined here as the number of pairs of nucleotides that are matching (are the same as each other) in a comparison of two aligned sequences divided by the total number of potential matches. "Maximum simple homology" is obtained when two sequences are aligned with each other so as to have the maximum number of paired matching nucleotides. In any event, the set of new sequences so obtained was referred to as the "candidate sequences". One of the candidate sequences was arbitrarily chosen and referred to as sequence 101. All the candidate sequences were checked against sequence 101, and sequences that contained 15 or more non-consecutive matches (i.e., a maximum simple homology of 15/24 (62.5%) or more were

eliminated. This results in a smaller set of candidate sequences from which another sequence is selected that is now referred to as sequence 102. The smaller set of candidate sequences is now compared to sequence 102 eliminating sequences that contained 15 or more non-consecutive matches and the process is repeated until there are no candidate sequences remaining. Also, any sequence selected from the candidate sequences is eliminated if it has 13 or more consecutive matches with any other previously selected candidate sequence.

Please replace the paragraph beginning at lines 15 and 23 on page 55 with the following amended paragraphs, respectively:

The additional set of 73 tag sequences so obtained (SEQ ID NOs:~~101 to 173~~ 1273 to 1345 of Table 1) is composed of sequences that when compared to any of SEQ ID NOs:~~1 to 100~~ 1173 to 1272 of Table I have no greater similarity than the sequences of the original 100 sequence tags of Table I. The sequence set as derived from the original family of non cross hybridizing sequences, SEQ ID NOs:~~1 to 173~~ 1173 to 1345 of Table 1, are expected to behave with similar hybridization properties to the sequences having SEQ ID NOs:~~1 to 100~~ 1173 to 1272 since it is understood that sequence similarity correlates directly with cross hybridization (Southern et al., Nat. Genet.; 21, 5-9: 1999).

The set of 173 24mer oligonucleotides were expanded to include those having SEQ ID NOs:~~174 to 210~~ 1346 to 1382 as follows. The 4mers WXYW, XYXW, WXXW, WYYW, XYYX, YXYX, YXXX and XYXY where W=G, X=A, and Y=U/T were used in combination with the fourteen 4mers used in the generation of SEQ ID NOs:~~1 to 173~~ 1173 to 1345 to generate potential 24-base oligonucleotides. Excluded from the set were those containing the sequence patterns GG, AAAA and TTTT. To be included in the set of additional 24mers, a sequence also had to have at least one of the 4mers containing two G's: WXYW (GATG), WYXW (GTAG), WXXW (GAAG), WYYW (GTTG) while also containing exactly six G's. Also

required for a 24mer to be included was that there be at most six bases between every neighboring pair of G's. Another way of putting this is that there are at most six non-G's between any two G's. Also, each G nearest the 5'-end of its oligonucleotide (the left-hand side as written in Table I) was required to occupy one of the first to seventh positions (counting the 5'- terminal position as the first position.) A set of candidate sequences was obtained by eliminating any new sequence that was found to have a maximum simple homology of 16/24 or more with any of the previous set of 173 oligonucleotides (Table 1, SEQ ID NOS:~~1 to 173~~ 1173 to 1345). As above, an arbitrary 174th sequence was chosen and candidate sequences eliminated by comparison therewith. In this case the permitted maximum degree of simple homology was 16/24. A second sequence was also eliminated if there were ten consecutive matches between the two (i.e., it was notionally possible to generate a phantom sequence containing a sequence of 10 bases that is identical to a sequence in each of the sequences being compared). A second sequence was also eliminated if it was possible to generate a phantom sequence 20 bases in length or greater.

Please replace the paragraphs beginning at lines 1 and 4 on page 57 with the following amended paragraphs, respectively:

The selection of sequences using this approach would be amenable to a computerized process. Thus for example, a string of 10 contiguous bases of the first 24mer of Table I could be selected: GATTTGTATTGATTGAGATTAAAG (SEQ ID NO:1173).

A string of contiguous bases from the second 24mer could then be selected and compared for maximum homology against the first chosen sequence: TGATTGTAGTATGTATTGATAAAG (SEQ ID NO:1174).

Please replace the paragraph beginning at line 1 on page 64 with the following amended paragraph:

A practical example of the aforementioned description is as follows:

For exon 1 of the human p53 tumor suppressor gene sequence the following tag-Reverse primer (SEQ ID NO:1171) was generated:

222087

222063

5'-PO4-GATTGTAAGATTTGATAAAAGTGT-A-TCCAGGGAAAGCGTGTCAACCGTCGT-3'

Tag Sequence # 3

Exon 1 Reverse

The numbering above the Exon-1 reverse primer represents the genomic nucleotide positions of the indicated bases.

The corresponding Exon-1 Forward primer sequence (SEQ ID NO:1172) is as follows:

221873

221896

5' -Biotin-TCATGGCGACTGTCCAGCTTG-3'

Please replace Table 1 beginning on page 87 with the following amended Table 1.

SEQ ID NO(1)	Sequence	No Assigned in Example 3	
1	1173	GATTGTATTGATTGAGATTAAAG	1
2	1174	TGATTGTAGTATGTATTGATAAAAG	2
3	1175	GATTGTAAGATTTGATAAAAGTGT	3
4	1176	GATTGAAAGATTATTGTAATGT	4
5	1177	GATTGATTATTGTGATTGAATTG	5
6	1178	GATTGATTGTAAAAGATTGTTGA	6
7	1179	ATGGTAAATTGTTAAATGAATTG	7
8	1180	ATGGATTGATAAAAGTAAATGA	
9	1181	GTAAGTAATGAATGTAAAAGGATT	8
10	1182	GATTGATTGATTGATTGATTGAT	
11	1183	TGATGATTAAAGAAAGTGTATTGAT	
12	1184	AAAGGATTGATTGATAAAAGTGT	
13	1185	TGTAGATTGTATGTATGTATGAT	10
14	1186	GATTGATAAAAGAAAGGATTGATT	
15	1187	GATTAAAGTGTATTGATGATTGTA	11
16	1188	AAAGAAAGAAAAGAAAAGTGT	12
17	1189	TGAAAAGGATTGATTGTATGTA	
18	1190	AAAGTGTAGATTGATTAAAGAAAG	
19	1191	AAAGTTGATTGATTGAAAAGGTAT	
20	1192	TTGATTGAGATTGATTGAGTAT	
21	1193	TGAATTGATGAATGAATGAAGTAT	15
22	1194	GTAATGAAGTATGTATGTAAGTAA	

SEQ ID NO (1)	Sequence	No Assigned in Example 3
23	TGATGATTGAATGAAGATTGATT	16
24	TGATAAAGTGATAAAAGGATTAAAG	17
25	TGATTTGAGTATTTGAGATTTGA	18
26	TGTAGTAAGATTGATTAAAGGTAA	
27	GTATAAAGGATTGATTGAAAAG	
28	GTATTTGAGTAAGTAATTGATTGA	19
29	GTAAAAAGTTGAGTATTGAAAAAG	
30	GATTTGATAAAGGATTGTATTGA	
31	GATTGTATTGAAGTATTGTAAAAG	20
32	TGATGATTTGATGAAAAAGTTGA	
33	TGATTTGAGATTAAAGAAAGGATT	21
34	TGATTGAATTGAGTAAAAGGATT	22
35	AAAGTGTAAAAGGATTGATGTAT	
36	AAAGGTATTGAGATTGATTGAA	
37	AAAGTTGAGATTGAAATGATTGAA	23
38	TGTATTGAAAAGGTATGATTGAA	
39	GTATTGTATTGAAAAGTAATTGA	24
40	TTGAGTAATGATAAAGTGAAGATT	
41	TGAAGATTGAAAGTAATTGAAAAG	25
42	TGAAAAAGTGTAGATTGAGTAA	26
43	TGTATGAATGAAGATTGATTGTA	
44	AAAGTTGAGTATTGATTGAAAAG	27
45	GATTTGTAGATTGTATTGAGATT	
46	AAAGAAAGGATTGTTGAGTAAAGATT	29
47	GTAAAAAGAAAGGTATAAAGGTAA	30
48	GATTAAAGTTGATTGAAAAGTGA	31
49	TGAAAAAGGTAAATTGATGTATGAA	
50	AAAGGATTAAAGTGAAGTAATTGA	33
51	ATGAATTGGTATGTATATGAATGA	34
52	TGAAATGAATGAATGATGAAATTG	35
53	ATTGATTGTGAATGAAATGAATTG	36
54	ATTGAAAGATGAAAAGATGAAAAG	37
55	ATTGTTGAAAAGTGTAAATGATTGA	38
56	ATGATGTAATGAAAAGATTGTGTA	39
57	AAAGATTGAAAGATGATGTAATTG	
58	ATTGATGAGTATATTGTTGAGTAA	41
59	AAAGATTGTTGAAATTGATGATGAA	
60	AAAGGTATATTGTGTAATGAGTAA	
61	TGTAATGAGTATTGTAATTGAAAG	43
62	GTATAAAGAAAGATTGGTAAATGA	44
63	TTGAGTAATTGAAATTGTTGATGTA	45
64	TGTATTGAATGAAATTGTTGATGTA	46
65	TGTAATTGGTAAATGAGTAAAAG	
66	TGAATGAAATTGATGAGTATAAAG	
67	GTAAGTAAATTGAAAGATTGATGA	49
68	GTAATGATGATATTGGTATATTG	50
69	ATTGTTGATGATTGATTGAAATGA	51
70	ATTGTGAAGTATAAAGATGATTGA	52
71	ATGAAAAGTTGAGTAAATTGTTGAT	
72	ATGAATTGAAAGTGTAGTAAAG	54
73	GTAATTGATGAAAAGTTGATGAT	
74	AAAGTGTATATGAGTAAATTG	56
75	GTAATGATAAAGATGATGATATTG	57

SEQ ID NO (1)	Sequence	No Assigned in Example 3
76	TTGAAAAGATTGTAATGATATGA	
77	AAAGTGAAAAAGATTGATTGATGA	59
78	ATTGATGAGATTGATTATTGTGTA	
79	ATGAGATTATTGGATTGGTAGATT	60
80	TGAAGATTATGAATTGTAAGATT	61
81	ATTGGATTATGAGATTATGATTGA	62
82	ATTGTTGAATTGGATTAAGATGA	
83	AAAGATGAGTAAGTAAATTGGATT	
84	AAAGGTAAGATTATTGATGAAAAG	65
85	ATTGATGAGATTAAAGTTGAATTG	
86	GATTATTGGATTATGAAAAGGATT	
87	GATTGTAATTGTTGAGTAAATGA	67
88	AAAGAAAGATTGTTGAGATTATGA	68
89	GTATAAAGGATTGGATTGAAATTGATGA	
90	TTGAGATTGTAATGAAATTGTTGA	
91	GTATATTGATTGTGTAATGAAAAG	
92	TGATATGAAATTGGATTATTGGTAT	70
93	ATGAATGATGAATGATGATTATTG	
94	ATGAATTGATTGGATTGTAATGAT	71
95	GATTGTAATTGAGTAAATTGATGA	
96	GATTATTGGATTAAGGTTAAATGAA	72
97	ATTGTTGAATTGATGAGATTGAT	73
98	GATTATGAGTAAATTGATTGTGAT	
99	GATTATTGTTGATGAATGATATTG	
100	TGTAAAAGATTGAAAGGTATGATT	75
101	GTATTTAGATGAGTTGTTAGATT	76
102	TGAAGTTATGTAATAGAAAAGTGT	
103	GTATGTATTGTATGTTAGTTAATTG	77
104	TGATATAGATAGTTAGATAGATAG	78
105	ATGATGATGTATTGTTAGTTATGAA	79
106	TTAGTGAATGTATTAGTTGATGTA	
107	GTTAGTTAGATTATTGTTAGTTAG	80
108	GTTAATTGTTAGTTGTTATTGAA	
109	GTTATGAAATAGTGTATTGTTAG	
110	ATTGTTAGAAAGTGTAGATTAAAG	81
111	ATGAGTATGTTATTAGTGTATGTA	82
112	TGTAATAGTGAAGTTAGATTGTAT	83
113	ATTGATAGATGATTAGTTAGTTGA	84
114	ATGAGTTGTTATGAGATTAAAG	
115	TGATGTTGATTATGATGTTAGTAT	85
116	ATGAGTTAGTTATGAAATTAGATGA	
117	ATTGTTAGTGTAGTTAGTAATTAG	86
118	TGATGTAAGTATTGATGTTAGTT	87
119	GATTGTAATAGAAAGTGAAGTAA	88
120	ATTGTTGATGAAAGTATTGTATGAT	
121	ATAGTGTATTGAAAGATTGTTA	
122	TTAGATGAAATTGTAAGTATTAG	90
123	GTAAGTTATGATTGATGTTATGAA	91
124	GTATTGATGTTAAAGTGTAAATAG	92
125	GATTGTAAGTAAGATTGTATATTG	
126	GTTTGATTGTAATAGTGTATTGAA	93
127	GTTTGATTGTAATAGTGTATTGAA	
128	TGTATGTTAGTATTAGAAAGATGA	
1300		

SEQ	ID NO(1)	Sequence	No Assigned in Example 3
129	1301	ATGAATTGTGATAAAGAAAGTTAG	
130	1302	TTAGTGTAGTAAGTTAAAGTGTA	95
131	1303	GTATGATTGTTGTAATTAGTGAT	
132	1304	GTTAAAGTTAGTTGAGTTAGTAT	96
133	1305	ATAGTGTATGTAGATTATGAGATT	97
134	1306	TTGAATGATTAGTTGAGTATGATT	98
135	1307	GTATGTAAGTTAGTATGATTGAA	
136	1308	TGTAGTATATTGTTGAATTGTCAT	
137	1309	ATAGTATTGTATGTATGATAAAG	
138	1310	TTAGTGATTGATGTATATTGAAAG	
139	1311	GTAAGATTATGAGTTATGATGTA	
140	1312	GTATGAAATTGTTAGTGTAGATT	99
141	1313	GTAGATTGAAATGATGTAGATT	100
142	1314	TTAGTGATTGAAATGATGTAGATT	
143	1315	AAAGTGTAGTTATTAGTTAGTTAG	
144	1316	AAAGAAAAGTGTATGATGTTATTAG	
145	1317	GATTGTATATTGTGTATGATGATT	
146	1318	TTGAGATTGTTATGATATGAGTAT	
147	1319	ATGAGTATGATTGTTATGATGTTT	
148	1320	TGATTTAGTGAATTGTTATTAG	
149	1321	TGAATGTATGTAGTATGTTGTTA	
150	1322	GTTAGTATTGATGATTATGAGTTA	
151	1323	GTATATTGTGATTTAGTTGAGATT	
152	1324	GTTAGTTAAAGTTGAGATTGTTT	
153	1325	GTATATTGTTAGATGAGATTGTA	
154	1326	TGATGTATGTTAGTTATGAATGA	
155	1327	TGAGTATGTAATGTAGTATTGAA	
156	1328	ATGAGTTATGATTGAGTTAGTAT	
157	1329	TGTATGATGATTATAGTTGAGTAA	
158	1330	ATTGATGAATGAGTTGTATAAAG	
159	1331	TTGAGTTATGATTAGAAAGAAAG	
160	1332	TGATATTGATGAGTTAGTATTGAA	
161	1333	ATAGAAAAGTGAATGAGTATGTTA	
162	1334	TTGATGTAGATTGATGTATATAG	
163	1335	TTGAGATTAGTGTAGTTATAG	
164	1336	TGATGTTAGATTGTTGATTATTG	
165	1337	TGTATTAGATAGTGTAGTTGAAATGA	
166	1338	GATTATGATGAATGTAGTATGTA	
167	1339	TGAATGATTGATATGAATAGTGT	
168	1340	GTAATGATTGATGTATTGAGTTT	
169	1341	TGAGTATGATTGATGATAAAG	
170	1342	TGAAGATTGTTATTAGTGTAT	
171	1343	GTATTTGAATGATGTAATAGTGT	
172	1344	GTATATGATGTATTAGATTGAAAG	
173	1345	AAAGTTAGATTGAAAGTGTATAAAG	
174	1346	GTAAGATGTTGATATAGAAGATTA	9
175	1347	TAATATGAGATGAAAGTGAATTAG	
176	1348	TTAGTGAAGAAGTATAGTTATTG	13
177	1349	GTAGTTGAGAAGATAGTAAATTAA	
178	1350	ATGAGATGATATTGAGAAGTAAT	
179	1351	GATGTGAAGAAGATGAATATATAT	
180	1352	AAAGTATAGTAAGATGTATAGTAG	14
181	1353	GAAGTAATATGAGTAGTTGAATAT	

SEQ ID NO (1)	Sequence	No Assigned in Example 3
182	TTGATAATGTTGTTGTTGTAG	28
183	TGAAGAAGAAAGTATAATGATGAA	
184	GTAGATTAGTTGAAGTGAAATAAT	32
185	TATAGTAGTGAAGATGATATATGA	
186	TATAATGAGTTGTTAGATATGTTG	
187	GTGTGAAATTAGATGTGAAATAT	
188	TAATGTTGTGAATAATGTAGAAAG	40
189	GTTTATAGTGAATATGAAGATAG	42
190	ATTATGAAGTAAGTTAATGAGAAAG	47
191	GATGAAAGTAATGTTTATTGTGAA	
192	ATTATTGAGATGTGAAGTTGTTT	48
193	TGAGAAGATGAGATGTATAATTAA	53
194	TAATTTGAGTTGTGTATATAGTAG	
195	TGATATTAGTAAGAAGTTGAATAG	
196	GTTAGTTATTGAGAAGTGTATATA	55
197	GTAGTAATGTTAATGAATTAGTAG	58
198	GTTTGTGTTGATGTGATTGAATAAT	
199	GTAAGTAGTAATTTGAATATGTAG	64
200	GTTTGAAGATATGTTGAAGTATA	
201	ATGATAATTGAAGATGTAATGTTG	
202	GTAGATAGTATAGTTGTAATGTTA	66
203	GATGTGAATGTAATATGTTTATAG	69
204	TGAAATTAGTTGTAAGATGTGTA	74
205	TGTAGTATAAAGTATATGAAGTAG	63
206	ATATGTTGAGTTGATAGTATA	89
207	ATTATTGAGTAGAAAGATAGAAAG	94
208	GTGTTGAATATTGAATATAGTTG	
209	ATGAGAAGTTAGTAATGTAAATAG	
210	TGAAATGAGAAGATTAATGAGTTT	

Please insert the sequence listing provided on pages 1/368 to 368/368 in the accompanying Response to the Notice to File Missing Parts of June 21, 2004, and submitted under separate cover, into the disclosure.